

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-204419

(43) 公開日 平成9年(1997)8月5日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	P 1	技術表示箇所
G 0 6 F 17/21			G 0 6 F 15/20	5 3 0 A
17/60				5 4 7 G
19/00			15/21	Z
G 0 6 T 11/60			15/22	C
			15/62	3 2 5 P
審査請求 未請求 請求項の数28 O L (全 18 頁)				

(21) 出願番号 特願平8-10045

(22) 出願日 平成8年(1996)1月24日

(71) 出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72) 発明者 齋藤 和之

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

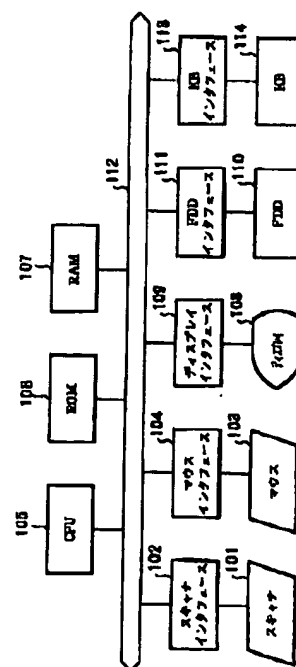
(74) 代理人 弁理士 大塚 康徳 (外1名)

(54) 【発明の名称】 画像処理装置及びその方法

(57) 【要約】

【課題】 領域分割処理で生じる処理ミスの有無を容易に判断することができ、かつ、処理ミスがある場合にその処理ミスを容易に修正することができる画像処理装置及びその方法を提供する。

【解決手段】 複数種類の戸籍文書に関するそれぞれのフォーマット情報をROM106に記憶する。そして、戸籍文書に含まれる罫線を抽出し、抽出される罫線の特徴に基づいて、戸籍文書の種類を識別する。識別された種類に対応する戸籍文書のフォーマット情報に基づいて、抽出結果の正否を判定する。更に、判定結果に基づいて、報知する。



(2)

特開平9-204419

1

【特許請求の範囲】

【請求項1】 所定のフォーマットを有する原稿画像に含まれる罫線を抽出する画像処理装置であって、前記所定のフォーマットに関するフォーマット情報を記憶する記憶手段と、

前記所定のフォーマットを有する原稿画像に含まれる罫線を抽出する抽出手段と、

前記フォーマット情報に基づいて、前記抽出手段による抽出結果の正否を判定する判定手段と、

前記判定手段の判定結果に基づいて、報知する報知手段とを備えることを特徴とする画像処理装置。

【請求項2】 所定のフォーマットを有する原稿画像に含まれる罫線を抽出する画像処理装置であって、複数種類の所定のフォーマットを有する原稿画像に関するそれぞれのフォーマット情報を記憶する記憶手段と、前記所定のフォーマットを有する原稿画像に含まれる罫線を抽出する抽出手段と、

前記抽出手段で抽出される罫線の特徴に基づいて、前記所定のフォーマットを有する原稿画像の種類を識別する識別手段と、

前記識別手段で識別された種類に対応する所定のフォーマットを有する原稿画像のフォーマット情報に基づいて、前記抽出手段による抽出結果の正否を判定する判定手段と、

前記判定手段の判定結果に基づいて、報知する報知手段とを備えることを特徴とする画像処理装置。

【請求項3】 前記判定手段の判定の結果、前記抽出手段による抽出結果が正である場合、該抽出結果に基づいて、前記所定のフォーマットを有する原稿画像を複数の領域に分割する分割手段を更に備えることを特徴とする請求項1又は請求項2に記載の画像処理装置。

【請求項4】 前記判定手段の判定の結果、前記抽出手段による抽出結果に否がある場合、該抽出結果の否を修正する修正手段を更に備えることを特徴とする請求項1又は請求項2に記載の画像処理装置。

【請求項5】 前記修正手段による修正を行うか否かを指示する指示手段を更に備えることを特徴とする請求項4に記載の画像処理装置。

【請求項6】 前記分割手段は、前記修正手段で修正された抽出結果に基づいて、前記所定のフォーマットを有する原稿画像を複数の領域に分割することを特徴とする請求項5に記載の画像処理装置。

【請求項7】 前記報知手段は、前記判定手段による判定結果を表示することを特徴とする請求項1又は請求項2に記載の画像処理装置。

【請求項8】 前記報知手段は、前記判定手段による判定結果を音を用いて報知することを特徴とする請求項1又は請求項2に記載の画像処理装置。

【請求項9】 前記判定手段の判定の結果、前記抽出手段による抽出結果に否がある場合、抽出結果に否が生じ

2

た可能性のある領域を、抽出結果に否が生じていない領域と区別して表示する表示手段を更に備えることを特徴とする請求項1又は請求項2に記載の画像処理装置。

【請求項10】 前記フォーマット情報に基づいて、前記分割手段による分割結果の正否を判定する判定手段を更に備えることを特徴とする請求項1又は請求項2に記載の画像処理装置。

【請求項11】 前記判定手段の判定の結果、前記分割手段による分割結果に否がある場合、該分割結果の否を修正する修正手段を更に備えることを特徴とする請求項10に記載の画像処理装置。

【請求項12】 前記判定手段の判定の結果、前記分割手段による分割結果の正否を表示する表示手段を更に備えることを特徴とする請求項10に記載の画像処理装置。

【請求項13】 前記判定手段の判定の結果、前記分割手段による分割結果に否がある場合、その旨を報知する報知手段を更に備えることを特徴とする請求項10に記載の画像処理装置。

20 【請求項14】 所定のフォーマットを有する原稿画像に含まれる罫線を抽出する画像処理方法であって、前記所定のフォーマットに関するフォーマット情報を記憶媒体に記憶する記憶工程と、前記所定のフォーマットを有する原稿画像に含まれる罫線を抽出する抽出工程と、前記フォーマット情報に基づいて、前記抽出工程による抽出結果の正否を判定する判定工程と、前記判定手段の判定結果に基づいて、報知する報知工程とを備えることを特徴とする画像処理方法。

30 【請求項15】 所定のフォーマットを有する原稿画像に含まれる罫線を抽出する画像処理方法であって、複数種類の所定のフォーマットを有する原稿画像に関するそれぞれのフォーマット情報を記憶媒体に記憶する記憶工程と、前記所定のフォーマットを有する原稿画像に含まれる罫線を抽出する抽出工程と、前記抽出工程で抽出される罫線の特徴に基づいて、前記所定のフォーマットを有する原稿画像の種類を識別する識別工程と、

40 前記識別工程で識別された種類に対応する所定のフォーマットを有する原稿画像のフォーマット情報に基づいて、前記抽出工程による抽出結果の正否を判定する判定工程と、前記判定工程の判定結果に基づいて、報知する報知工程とを備えることを特徴とする画像処理方法。

【請求項16】 前記判定工程の判定の結果、前記抽出工程による抽出結果が正である場合、該抽出結果に基づいて、前記所定のフォーマットを有する原稿画像を複数の領域に分割する分割工程を更に備えることを特徴とする請求項14又は請求項15に記載の画像処理方法。

50

3

【請求項 17】 前記判定工程の判定の結果、前記抽出工程による抽出結果に否がある場合、該抽出結果の否を修正する修正工程を更に備えることを特徴とする請求項 14 又は請求項 15 に記載の画像処理方法。

【請求項 18】 前記修正工程による修正を行うか否かを指示する指示工程を更に備えることを特徴とする請求項 17 に記載の画像処理方法。

【請求項 19】 前記分割工程は、前記修正工程で修正された抽出結果に基づいて、前記所定のフォーマットを有する原稿画像を複数の領域に分割することを特徴とする請求項 18 に記載の画像処理方法。

【請求項 20】 前記報知工程は、前記判定工程による判定結果を表示することを特徴とする請求項 14 又は請求項 15 に記載の画像処理方法。

【請求項 21】 前記報知工程は、前記判定工程による判定結果を音を用いて報知することを特徴とする請求項 14 又は請求項 15 に記載の画像処理方法。

【請求項 22】 前記判定工程の判定の結果、前記抽出工程による抽出結果に否がある場合、抽出結果に否が生じた可能性のある領域を、抽出結果に否が生じていない領域と区別して表示する表示工程を更に備えることを特徴とする請求項 14 又は請求項 15 に記載の画像処理方法。

【請求項 23】 前記フォーマット情報に基づいて、前記分割工程による分割結果の正否を判定する判定工程を更に備えることを特徴とする請求項 14 又は請求項 15 に記載の画像処理方法。

【請求項 24】 前記判定工程の判定の結果、前記分割工程による分割結果に否がある場合、該分割結果の否を修正する修正工程を更に備えることを特徴とする請求項 23 に記載の画像処理方法。

【請求項 25】 前記判定工程の判定の結果、前記分割工程による分割結果の正否を表示する表示工程を更に備えることを特徴とする請求項 23 に記載の画像処理方法。

【請求項 26】 前記判定工程の判定の結果、前記分割工程による分割結果に否がある場合、その旨を報知する報知工程を更に備えることを特徴とする請求項 23 に記載の画像処理方法。

【請求項 27】 画像処理のプログラムコードが格納されたコンピュータ可読メモリであって、前記所定のフォーマットに関するフォーマット情報を記憶媒体に記憶する記憶工程のコードと、前記所定のフォーマットを有する原稿画像に含まれる罫線を抽出する抽出工程のコードと、前記フォーマット情報に基づいて、前記抽出工程による抽出結果の正否を判定する判定工程のコードと、前記判定手段の判定結果に基づいて、報知する報知工程のコードとを備えることを特徴とするコンピュータ可読メモリ。

(3)

特開平 9-204419

4

【請求項 28】 画像処理のプログラムコードが格納されたコンピュータ可読メモリであって、

複数種類の所定のフォーマットを有する原稿画像に関するそれぞれのフォーマット情報を記憶媒体に記憶する記憶工程のコードと、

前記所定のフォーマットを有する原稿画像に含まれる罫線を抽出する抽出工程のコードと、

前記抽出工程で抽出される罫線の特徴に基づいて、前記所定のフォーマットを有する原稿画像の種類を識別する識別工程のコードと、

前記識別工程で識別された種類に対応する所定のフォーマットを有する原稿画像のフォーマット情報に基づいて、前記抽出工程による抽出結果の正否を判定する判定工程のコードと、

前記判定工程の判定結果に基づいて、報知する報知工程のコードとを備えることを特徴とするコンピュータ可読メモリ。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、所定のフォーマットを有する原稿画像に含まれる罫線を抽出する画像処理装置及びその方法に関するものである。

【0002】

【従来の技術】従来の画像処理装置においては複数の罫線を有する原稿画像に対し、その原稿画像に含まれる罫線毎に領域を分割する領域分割処理によって、罫線に囲まれた領域を属性の一つとしてその領域を抽出していた。

【0003】

【発明が解決しようとする課題】しかしながら、上記従来の画像処理装置では、領域分割処理によって罫線に囲まれた領域を抽出できても、領域分割処理のミスの有無はオペレータが目視で確認する必要があった。また、原稿画像に含まれる所望の罫線に囲まれた領域を、その原稿画像を構成する項目の一つとして抽出しようとしても領域分割処理にミスが有る場合は、その項目を誤って取り出してしまうという問題点があった。

【0004】本発明は上記の問題点を鑑みてなされたものであり、領域分割処理で生じる処理ミスの有無を容易に判断することができ、かつ、処理ミスがある場合にその処理ミスを容易に修正することができる画像処理装置及びその方法を提供することを目的としている。また、原稿画像に含まれる罫線を抽出する罫線抽出処理で生じる処理ミスを修正することで、常に正確な罫線を抽出が可能とし、更に、抽出された罫線に基づいて、領域分割処理を行うことで、領域分割処理で生じる処理ミスを軽減する画像処理装置及びその方法を提供することを目的としている。

【0005】また、領域分割処理における操作の作業効率を向上する画像処理装置及びその方法を提供すること

5

を目的としている。

【0006】

【課題を解決するための手段】上記の目的を達成するため本発明による画像処理装置は以下の構成を備える。即ち、所定のフォーマットを有する原稿画像に含まれる罫線を抽出する画像処理装置であって、前記所定のフォーマットに関するフォーマット情報を記憶する記憶手段と、前記所定のフォーマットを有する原稿画像に含まれる罫線を抽出する抽出手段と、前記フォーマット情報に基づいて、前記抽出手段による抽出結果の正否を判定する判定手段と、前記判定手段の判定結果に基づいて、報知する報知手段とを備える。

【0007】上記の他の目的を達成するための本発明による画像処理装置は以下の構成を備える。即ち、所定のフォーマットを有する原稿画像に含まれる罫線を抽出する画像処理装置であって、複数種類の所定のフォーマットを有する原稿画像に関するそれぞれのフォーマット情報を記憶する記憶手段と、前記所定のフォーマットを有する原稿画像に含まれる罫線を抽出する抽出手段と、前記抽出手段で抽出される罫線の特徴に基づいて、前記所定のフォーマットを有する原稿画像の種類を識別する識別手段と、前記識別手段で識別された種類に対応する所定のフォーマットを有する原稿画像のフォーマット情報に基づいて、前記抽出手段による抽出結果の正否を判定する判定手段と、前記判定手段の判定結果に基づいて、報知する報知手段とを備える。

【0008】また、好ましくは、前記判定手段の判定の結果、前記抽出手段による抽出結果が正である場合、該抽出結果に基づいて、前記所定のフォーマットを有する原稿画像を複数の領域に分割する分割手段を更に備える。また、好ましくは、前記判定手段の判定の結果、前記抽出手段による抽出結果に否がある場合、該抽出結果の否を修正する修正手段を更に備える。修正手段を備えることで、抽出結果の否を修正することができるからである。

【0009】また、好ましくは、前記修正手段による修正を行うか否かを指示する指示手段を更に備える。また、好ましくは、前記分割手段は、前記修正手段で修正された抽出結果に基づいて、前記所定のフォーマットを有する原稿画像を複数の領域に分割する。

【0010】また、好ましくは、前記報知手段は、前記判定手段による判定結果を表示する。また、好ましくは、前記報知手段は、前記判定手段による判定結果を音を用いて報知する。また、好ましくは、前記判定手段の判定の結果、前記抽出手段による抽出結果に否がある場合、抽出結果に否が生じた可能性のある領域を、抽出結果に否が生じていない領域と区別して表示する表示手段を更に備える。区別して表示することで、ユーザに抽出結果に否が生じた可能性のある罫線を把握しやすくなることができるからである。

(4)

特開平9-204419

6

【0011】また、好ましくは、前記フォーマット情報に基づいて、前記分割手段による分割結果の正否を判定する判定手段を更に備える。また、好ましくは、前記判定手段の判定の結果、前記分割手段による分割結果に否がある場合、該分割結果の否を修正する修正手段を更に備える。修正手段を備えることで、分割結果の否を修正することができるからである。

【0012】また、好ましくは、前記判定手段の判定の結果、前記分割手段による分割結果の正否を表示する表示手段を更に備える。表示手段を備えることで、ユーザに分割結果の正否を知らせることができるからである。また、好ましくは、前記判定手段の判定の結果、前記分割手段による分割結果に否がある場合、その旨を報知する報知手段を更に備える。報知手段を備えることで、ユーザに分割結果に否があったことを知らせることができるからである。

【0013】上記の目的を達成するための本発明による画像処理方法は以下の構成を備える。即ち、所定のフォーマットを有する原稿画像に含まれる罫線を抽出する画像処理方法であって、前記所定のフォーマットに関するフォーマット情報を記憶媒体に記憶する記憶工程と、前記所定のフォーマットを有する原稿画像に含まれる罫線を抽出する抽出工程と、前記フォーマット情報に基づいて、前記抽出工程による抽出結果の正否を判定する判定工程と、前記判定手段の判定結果に基づいて、報知する報知工程とを備える。

【0014】上記の他の目的を達成するための本発明による画像処理方法は以下の構成を備える。即ち、所定のフォーマットを有する原稿画像に含まれる罫線を抽出する画像処理方法であって、複数種類の所定のフォーマットを有する原稿画像に関するそれぞれのフォーマット情報を記憶媒体に記憶する記憶工程と、前記所定のフォーマットを有する原稿画像に含まれる罫線を抽出する抽出工程と、前記抽出工程で抽出される罫線の特徴に基づいて、前記所定のフォーマットを有する原稿画像の種類を識別する識別工程と、前記識別工程で識別された種類に対応する所定のフォーマットを有する原稿画像のフォーマット情報に基づいて、前記抽出工程による抽出結果の正否を判定する判定工程と、前記判定工程の判定結果に基づいて、報知する報知工程とを備える。

【0015】また、好ましくは、前記判定工程の判定の結果、前記抽出工程による抽出結果が正である場合、該抽出結果に基づいて、前記所定のフォーマットを有する原稿画像を複数の領域に分割する分割工程を更に備える。また、好ましくは、前記判定工程の判定の結果、前記抽出工程による抽出結果に否がある場合、該抽出結果の否を修正する修正工程を更に備える。

【0016】また、好ましくは、前記修正工程による修正を行うか否かを指示する指示工程を更に備える。また、好ましくは、前記分割工程は、前記修正工程で修正

7

された抽出結果に基づいて、前記所定のフォーマットを有する原稿画像を複数の領域に分割する。

【0017】また、好ましくは、前記報知工程は、前記判定工程による判定結果を表示する。また、好ましくは、前記報知工程は、前記判定工程による判定結果を音を用いて報知する。また、好ましくは、前記判定工程の判定の結果、前記抽出工程による抽出結果に否がある場合、抽出結果に否が生じた可能性のある領域を、抽出結果に否が生じていない領域と区別して表示する表示工程を更に備える。

【0018】また、好ましくは、前記フォーマット情報に基づいて、前記分割工程による分割結果の正否を判定する判定工程を更に備える。また、好ましくは、前記判定工程の判定の結果、前記分割工程による分割結果に否がある場合、該分割結果の否を修正する修正工程を更に備える。

【0019】また、好ましくは、前記判定工程の判定の結果、前記分割工程による分割結果の正否を表示する表示工程を更に備える。また、好ましくは、前記判定工程の判定の結果、前記分割工程による分割結果に否がある場合、その旨を報知する報知工程を更に備える。

【0020】上記の目的を達成するための本発明によるコンピュータ可読メモリは以下の構成を備える。即ち、画像処理のプログラムコードが格納されたコンピュータ可読メモリであって、前記所定のフォーマットに関するフォーマット情報を記憶媒体に記憶する記憶工程のコードと、前記所定のフォーマットを有する原稿画像に含まれる罫線を抽出する抽出工程のコードと、前記フォーマット情報に基づいて、前記抽出工程による抽出結果の正否を判定する判定工程のコードと、前記判定手段の判定結果に基づいて、報知する報知工程のコードとを備える。

【0021】上記の他の目的を達成するための本発明によるコンピュータ可読メモリは以下の構成を備える。即ち、画像処理のプログラムコードが格納されたコンピュータ可読メモリであって、複数種類の所定のフォーマットを有する原稿画像に関するそれぞれのフォーマット情報を記憶媒体に記憶する記憶工程のコードと、前記所定のフォーマットを有する原稿画像に含まれる罫線を抽出する抽出工程のコードと、前記抽出工程で抽出される罫線の特徴に基づいて、前記所定のフォーマットを有する原稿画像の種類を識別する識別工程のコードと、前記識別工程で識別された種類に対応する所定のフォーマットを有する原稿画像のフォーマット情報に基づいて、前記抽出工程による抽出結果の正否を判定する判定工程のコードと、前記判定工程の判定結果に基づいて、報知する報知工程のコードとを備える。

【0022】

【発明の実施の形態】以下、図面を参照して本発明の好適な実施の形態を詳細に説明する。図1は本実施の形態

(5)

特開平9-204419

8

の画像処理装置の機能構成を示すブロック図である。図1において、101はスキャナであり、原稿画像に光を照射し、その反射光を読み取り電気信号に変換する。また、スキャナ101より読み込まれた原稿画像は、イメージデータとしてRAM107のVRAM領域に格納される。102はスキャナインターフェースであり、スキャナ101で得られた電気信号を2値のデジタル電気信号に変換し、他の装置構成要素に伝送する。103はポインティングデバイスとして機能するマウスであり、マウス103を移動することでディスプレイ108上に表示されるカーソル100を移動する。また、画面に表示されるメニューにカーソル100を位置し、そのメニュー上で付帯するボタン（不図示）を用いてドラッグ操作、クリック操作を行うことで、その内容を実行することができる。104はマウスインターフェースであり、マウス103からの信号を受け、それを他の装置構成要素に伝送する。114はKB（キーボード）であり、キーボードを介してデータの入力や命令の実行を行う。113はKBインターフェースであり、KB114からの信号を受け、それを他の装置構成要素に伝送する。

【0023】105はCPUであり、装置全体の制御及び後述する領域分割等の本発明で説明される処理や認識処理を実行する。106はROMであり、CPU105が実行する制御プログラム、各種処理プログラムやフォントデータ等の各種データを格納している。107はRAMであり、原稿画像の展開や領域分割処理のための作業領域等として用いる。また、スキャナ101より読み込まれたイメージデータ（2値画像データ）を格納するための所定の記憶領域としてVRAM領域を持つ。108はディスプレイであり、RAM107のVRAM領域に格納されているイメージデータや本実施の形態で実行される罫線抽出処理結果等を表示する。109はディスプレイインターフェースである。

【0024】110はFDD（フロッピーディスクドライブ）であり、FDを装着しデータの読み書きが可能である。また、装着されたFD（不図示）に後述する処理フローのプログラムを書き込み、該プログラムを本装置のRAM107に読み込むことで処理を実行させることができる。111はFDDインターフェースである。そして、112は各装置構成要素を接続するバスである。

【0025】次に本実施の形態で実行される処理について、図2のフローチャートを用いて説明する。図2は本実施の形態で実行される処理の処理フローを示すフローチャートである。尚、本実施の形態で実行される処理は、あらかじめ原稿画像を構成する罫線で囲まれる領域（以下、この領域を項目と呼ぶ）が決まっているような、例えば、図6に示すような戸籍文書等に適用した場合について説明するが、これに限定されるものではない。例えば、構成が決まっている原稿画像であるならば、どのようなものでも良い。

9

【0026】まず、ステップS201で、原稿画像をスキャナ101で読み込み2値画像データに変換する。ステップS202で、得られた2値画像データに対して、2値画像データに含まれる罫線を抽出する罫線抽出処理を行う。そして、抽出した罫線の座標データ等の情報をブロックデータとして図3のように記憶する。ここで、図3に示されるブロックデータについて説明する。

【0027】図3は本実施の形態のブロックデータの構成を示す図である。図3は、RAM107に展開される2値画像データに含まれる項目の位置を示すための始点座標(X1, Y1)、終点座標(X2, Y2)、有効フラグ、項目識別ID、認識順序、属性、組方向の情報を格納する。始点座標は、図6に示されるような各罫線の右端の位置座標、終点座標は、各罫線の左端の位置座標を示す。有効フラグは、抽出された罫線が有効であるか否かを示す。項目識別IDは、ここでは使用しない。認識順序も、ここでは使用しない。属性は、領域分割処理で分割単位である属性を格納する。本発明の場合、罫線を示す属性が格納される。組方向も、ここでは使用しない。

【0028】例えば、図6に示される戸籍文書に対し、上述のステップS202の罫線抽出処理を行うと、まず、戸籍文書に含まれる罫線を抽出し、その罫線の位置、長さ、本数等から戸籍文書の種別や戸籍文書を構成する項目の識別を行う。その罫線抽出処理の処理結果が、図7に示されるように、罫線701や罫線702や罫線703が存在すれば、その戸籍文書の項目の構成の種別(以下、「ページタイプ」と呼ぶ)は、筆頭者が記載されているページタイプ(以下、「筆頭者タイプ」と呼ぶ)と識別される。また、罫線701や罫線702が無く、罫線703のX1座標が上端の罫線704または下端の罫線705と同じであれば、筆頭者が記載されていないページタイプ(以下、「非筆頭者タイプ」と呼ぶ)と識別される。

【0029】更に、罫線703から下に連なる縦罫線の本数によって、戸籍文書の識別がなされる。例えば、図7に示される戸籍文書の場合は、養父および養母の両方が記載されているタイプであると識別される。次に、S203で、上述のステップS202の罫線抽出処理によって得られた罫線に基づいて、罫線抽出処理のミスの有無を判定する罫線抽出ミス判定処理を行う。このミスの判定は、上述のステップS202の罫線抽出処理によって得られた罫線と、予め罫線の構成が定まっている戸籍文書(図6参照)の罫線を比較することで罫線抽出処理のミスの有無を判定する。尚、予め罫線の構成が定まっている戸籍文書は、RAM106に格納されているとする。

【0030】罫線抽出処理のミスが有るとして判定されるのは、例えば、図8に示されるように、スキャナ101で読み取った原稿画像の罫線がかすれていて罫線80

(6)

特開平9-204419

10

1のように抽出できなかったり、あるいは罫線803のように抽出された罫線が短くなってしまった場合に、罫線抽出処理のミスがあると判定される。なぜなら、このまま、処理を続けてしまうと、後述するブロック抽出処理(ステップS207)の処理ミスを起こしたり、戸籍文書の種別や項目の識別が不可能となるからである。

【0031】ステップS204で、罫線抽出ミス判定処理の処理結果をディスプレイ108に表示する罫線抽出ミス判定結果表示処理を行う。ステップS205で、罫線抽出ミス判定結果表示処理の処理結果より、罫線抽出処理にミスが有るか否かを判定する。罫線抽出処理にミスがない場合、正常終了した旨をディスプレイ108に表示し、ステップS207に進む。一方、罫線抽出処理にミスがある場合、罫線抽出処理にミスがあった旨をディスプレイ108に表示し、罫線抽出処理のミスの修正をオペレータに促し、ステップS206に進む。

【0032】ステップS206で、罫線抽出ミス判定結果表示処理によって罫線抽出処理のミスがあった旨が表示された後、元原稿画像上に現在の罫線の罫線抽出状態が表示される。そして、オペレータは修正対象を拡大、縮小、移動、削除するか、新しいブロックを作成する等の画像処理によって罫線抽出処理の処理結果を修正する罫線修正処理を行う。

【0033】修正した結果は、オペレータがマウス103で修正終了ボタンを押す等によって指示することで、修正された項目に対応するブロックデータが更新された後、ステップS207に進む。例えば、罫線抽出ミス判定結果表示処理において、罫線抽出処理にミスがあった旨がディスプレイ108に表示されると、元原稿画像(図6の戸籍文書)の上に処理結果(図8)が表示される。

【0034】元原稿画像画像上に現在の処理結果(図8)が表示されると、オペレータはそこから、抽出されていない罫線801(図8)を図9に示すような新しい罫線901を作成する。同様にして、罫線803(図8)を修正する(不図示)。そして、オペレータがマウス103で修正終了のボタン902を指示することによって、修正された項目のブロックの内容は対応するブロックデータに更新される。

【0035】ステップS207で、原稿画像に含まれる罫線に囲まれた領域(項目)を抽出するブロック抽出処理を行う。そして、抽出した項目の座標データ等の情報をブロックデータとして図3のように記憶する。この場合、始点座標(X1, Y1)、終点座標(X2, Y2)、有効フラグ、項目識別ID、認識順序、属性、組方向の情報を格納する。始点座標は、図11に示されるような各項目の左上角の位置座標、終点座標は、各項目の左下角の位置座標を示す。有効フラグは、抽出された項目が有効であるか否かを示す。項目識別IDは、後述する設定方法によって各項目の項目識別IDを格納す

11

る。認識順序は、各項目に記入されている文字を認識する場合の認識順序を格納する。属性は、領域分割処理で分割単位である属性を格納する。本発明の場合、項目（罫線で囲まれた領域）を示す属性が格納される。組方向は、各項目に記入されている文字の方向（縦書きあるいは横書き）を示す情報を格納する。

【0036】次に、ステップS208で、原稿画像の項目の構成を示すフォーマット情報からその原稿画像の項目の構成を分類し、その原稿画像を構成する項目を識別する項目識別処理を行う。尚、原稿画像の項目の構成を識別する場合、その原稿画像を領域分割して得られるブロックの位置、形状、面積、数等によって、その原稿画像の項目の構成の種類を分類し、各項目を識別する。また、フォーマット情報はあらかじめROM106に記憶されているものとする。

【0037】また、本実施の形態の原稿画像として用いる戸籍文書の項目を分類し識別するために、各項目を分類し識別するためのID（項目識別ID）を図4の

(a)～(c)のように定義する。そして、各項目のIDを組み合わせることによって戸籍文書の項目を分類し識別する。ここで、図4の(a)～(c)に示されるID（項目識別ID）について説明する。

【0038】図4の(a)は、戸籍文書のページタイプの種別を定義するIDであり、戸籍文書の最右上にある罫線に囲まれたブロック領域（図6参照）に書かれている、「筆頭者」あるいは「非筆頭者」を種別として定義している。図4の(b)は、戸籍文書のページの記載者の種別を定義するIDであり、図4の(b)に示されるような種別のIDが定義されている。図4の(c)は、戸籍文書を構成する項目の項目名を定義するIDであり、図4の(c)に示されるような種別のIDが定義されている。

【0039】次に、ステップS208で実行される項目識別処理の詳細を、図10を用いて説明する。図10は本実施の形態の項目識別処理の詳細な処理フローを示すフローチャートである。まず、ステップS1001で、戸籍文書のページタイプの種別を分類し識別する。この識別の方法としては、戸籍文書の最右上にあるブロック（項目）の形状によって、その戸籍文書のページタイプを識別する。

【0040】尚、識別の方法としては、例えば、戸籍文書の最右上のブロックの幅がその高さよりも長い場合は「筆頭者タイプ」として識別し、それ以外は「非筆頭者タイプ」として識別することでページタイプを識別する。あるいは、最右上のブロックの幅が、そのブロックのY1座標と同じY1座標を持つブロックの幅よりも大きい場合は「筆頭者タイプ」として識別し、一方、最右上のブロックの幅が、そのブロックのY1座標と同じY1座標を持つブロックの幅が同じである場合は「非筆頭者タイプ」として識別することもできる。

(7)

特開平9-204419

12

【0041】再び、フローチャートの説明に戻る。ステップS1001において、戸籍文書の最右上のブロックの幅がその高さよりも長い場合は「筆頭者タイプ」として識別し、ステップS1004に進む。一方、戸籍文書の最右上のブロックの幅がその高さよりも短い場合は「非筆頭者タイプ」として識別し、ステップS1002に進む。

【0042】戸籍文書のページタイプが識別されたならば、次に、更に詳細な戸籍文書の識別と項目の識別の鍵となるブロック（以下、KEYブロックと呼ぶ）を、戸籍文書に構成される各項目の相対的な位置関係に検出する。まず、ページタイプが「筆頭者タイプ」の場合、ステップS1004で、「本籍」というタイトルがついているブロックと、編製された年月日が記載されている欄の最初の行のブロックと、入籍データ等が記載されている欄の最初の行のブロックと、「父」というタイトルがついているブロックをそれぞれKEY0～3ブロックとして検出し、各ブロックデータのアドレスを図5に示すようなKEYブロックデータ構造体に記憶する。

【0043】ここで図5に示すKEYブロックデータ構造体について説明する。図5は本実施の形態のKEYブロックデータ構造体の構成を示す図である。図5において、KEYブロック構造体は、戸籍文書を構成する項目の識別の鍵となる項目の情報を記憶するものであり、図に示すような8つのKEY0～5ブロックが存在する。そして、ページタイプの種別に応じて検出すべきKEYブロックが検出され、その検出されたKEYブロックのブロックデータを記憶する。

【0044】次に、ステップS1005で、検出したKEY0～3ブロックを用いて、各KEY0～3ブロックと共通する情報を持つ項目毎に、更に詳細に項目を抽出し、抽出した項目に対し、項目識別IDを設定する。この設定の詳細については、図11のフローチャートを用いて後述する。一方、ページタイプが「非筆頭者タイプ」の場合、ステップS1002で、1人目の入籍データが記載されている欄の最初の行のブロックと、1人目の「父」というタイトルがついているブロックと、2人目の入籍データ等が記載されている欄の最初の行のブロックと、2人目の「父」というタイトルがついているブロックをそれぞれKEY2～5ブロックとして検出し、その検出されたKEY2～5ブロックのブロックデータをKEYブロック構造体に記憶する。

【0045】次に、ステップS1003で、検出したKEY2～5ブロックを用いて、各KEY2～5ブロックと共通する情報を持つ項目毎に、更に詳細に項目を抽出し、抽出した項目に対し、項目識別IDを設定する。この設定の詳細については、図12のフローチャートを用いて後述する。まず、ページタイプが「筆頭者タイプ」の場合の項目識別IDを設定する処理について、図11のフローチャートを用いて説明する。

(8)

特開平9-204419

13

【0046】図11は本実施の形態のページタイプが「筆頭者タイプ」の場合の項目識別IDを設定する処理の処理フローを示すフローチャートである。まず、ステップS1101で、KEY0ブロック（図6の「本籍」というタイトルがついているブロック）の下にある3つの同じ幅を持つブロックに対して、それぞれ項目識別IDを設定する。ステップS1102で、KEY1ブロック（図6の編製事項が記載されている欄の最初の行のブロック）からKEY1ブロックを含めて左に4個のブロックおよびその下にあるブロックは、同じ項目群（編製事項の欄）とみなせるブロックであるので、これらのブロックに項目識別IDを設定する。

【0047】ステップS1103で、KEY2ブロック（図6の入籍データ等が記載されている欄の最初の行のブロック）からKEY2ブロックを含めて左に10個のブロックは、同じ項目群（入籍データ等の欄）とみなせるブロックであるので、これらのブロックに項目識別IDを設定する。ステップS1104で、KEY2ブロックを含む項目群の下にあるブロックは、KEY3ブロック（図6の「父」というタイトルがついているブロック）からKEY3ブロックを含めて左にあるKEY3ブロックと同じY1座標を持つブロックが存在し、そのブロック数を計数する。そして、ステップS1105で、計数されたブロック数に基づいて、KEY2ブロックを含む項目群の下にあるブロックの項目識別IDを設定する。この設定の詳細については、図13のフローチャートを用いて後述する。

【0048】次に、ページタイプが「非筆頭者タイプ」の場合の項目識別IDを設定する処理について、図12のフローチャートを用いて説明する。図12は本実施の形態のページタイプが「非筆頭者タイプ」の場合の項目識別IDを設定する処理の処理フローを示すフローチャートである。まず、ステップS1201で、KEY2ブロック（入籍データ等が記載されている欄の最初の行のブロック）からKEY2ブロックを含めて左に9個のブロックは、同じ項目群（入籍データ等の欄）とみなせるブロックであるので、これらのブロックに項目識別IDを設定する。

【0049】ステップS1202で、KEY4ブロック（2人目の入籍データ等が記載されている欄の最初の行のブロック）からKEY4ブロックを含めて左に9個のブロックは、同じ項目群（2人目の入籍データ等の欄）とみなせるブロックであるので、これらのブロックに項目識別IDを設定する。ステップS1203で、KEY2ブロックを含む項目群の下にあるブロックは、KEY3ブロック（「父」というタイトルがついているブロック）からKEY3ブロックを含めて左にあるKEY3ブロックと同じY1座標を持つブロックが存在し、そのブロック数を計数する。そして、ステップS1204で、計数されたブロック数に基づいて、KEY2ブロックを

14

含む項目群の下にあるブロックの項目識別IDを設定する。

【0050】ステップS1205で、KEY4ブロックを含む項目群の下にあるブロックは、KEY5ブロック（2人目の「父」というタイトルがついているブロック）からKEY5ブロックを含めて左にあるKEY5ブロックと同じY1座標を持つブロックが存在し、そのブロック数を計数する。そして、ステップS1206で、計数されたブロック数に基づいて、KEY4ブロックを含む項目群の下にあるブロックの項目識別IDを設定する。

【0051】次に、入籍データ等の欄の下にあるブロックの項目識別IDを設定する処理の詳細について、図13のフローチャートを用いて説明する。図13は本実施の形態の入籍データ等の欄の下にあるブロックの項目識別IDを設定する処理の処理フローを示すフローチャートである。尚、図13のフローチャートは、図11のフローチャートのステップS1104、図12のフローチャートのステップS1204、ステップS1205で実行される処理であるが、ここでは図11のフローチャートのステップS1104、つまり、ページタイプが「筆頭者タイプ」の戸籍文書である場合を、例に挙げて説明する。また、図12のフローチャートのステップS1204、ステップS1205、つまり、ページタイプが「非筆頭者タイプ」の戸籍文書である場合については、同様の処理で実現されるので、ここでは省略する。

【0052】まず、ステップS1301で、計数されたブロック数を判定し、ブロック数が4の場合は、実父母のみが記載されているタイプ（以下、「実子タイプ」と呼ぶ）と判定して、ステップS1302に進む。また、ブロック数が5の場合は、養父あるいは養母のどちらか一方のみが記載されているタイプ（以下、「養子タイプA」と呼ぶ）と判定して、ステップS1305に進む。更に、ブロック数が6の場合は、養父および養母の両方が記載されているタイプ（以下、「養子タイプB」と呼ぶ）と判定して、ステップS1309に進む。

【0053】計数されたブロック数が4の場合、つまり、「実子タイプ」である場合、ステップS1302で、KEY3ブロックから左に2番目のブロックの形状が横長であるか否かを判定する。横長である場合（ステップS1302でYES）、「既婚（夫か妻）タイプ」と判定して、ステップS1303に進む。一方、横長でない場合（ステップS1302でNO）、「未婚タイプ」と判定して、ステップS1304に進む。

【0054】ステップS1303では、「実子タイプ」及び「既婚（夫か妻）タイプ」として、KEY3ブロックとの相対位置関係から項目識別IDを設定する。また、ステップS1304では、「実子タイプ」及び「未婚タイプ」として、KEY3ブロックとの相対位置関係から項目識別IDを設定する。計数されたブロック数が

15

5の場合、つまり、「養子タイプA」である場合、ステップS1305で、KEY3ブロックから左に3番目のブロックの形状が横長であるか否かを判定する。横長である場合（ステップS1305でYES）、「既婚（夫か妻）タイプ」と判定して、ステップS1306に進む。一方、横長でない場合（ステップS1305でNO）、「未婚タイプ」と判定して、ステップS1307に進む。

【0055】ステップS1306では、「実子タイプ」及び「既婚（夫か妻）タイプ」として、KEY3ブロックとの相対位置関係から項目識別IDを設定する。また、ステップS1307では、「実子タイプ」及び「未婚タイプ」として、KEY3ブロックとの相対位置関係から項目識別IDを設定する。計数されたブロック数が6の場合、つまり、「養子タイプB」である場合、ステップS1309で、KEY3ブロックから左に4番目のブロックの形状が横長であるか否かを判定する。横長である場合（ステップS1309でYES）、「既婚（夫か妻）タイプ」と判定して、ステップS910に進む。一方、横長でない場合（ステップS1309でNO）、「未婚タイプ」と判定して、ステップS1311に進む。

【0056】ステップS1310では、「実子タイプ」及び「既婚（夫か妻）タイプ」として、KEY3ブロックとの相対位置関係から項目識別IDを設定する。また、ステップS1311では、「実子タイプ」及び「未婚タイプ」として、KEY3ブロックとの相対位置関係から項目識別IDを設定する。以上説明したように、本実施の形態によれば、戸籍文書に含まれる野線の構成を識別することで、野線抽出処理の処理ミスの有無を判定することができる。また、野線抽出処理の処理ミスがある場合でも、その処理ミスを修正することができる。

【0057】また、抽出された野線、あるいは修正済みの野線に基づいて、領域分割処理を行うので、領域分割処理を正確に行うことができる。更に、上述の野線抽出処理と、野線抽出ミス判定処理と、野線抽出ミス判定結果表示処理と、ブロック抽出処理と、項目識別処理を設けて野線抽出処理のミスの修正を行うことができるようにすることにより、ミスの有無の判断が容易になり、操作性を向上し、処理全体の時間の短縮が図ることができる。

【0058】以上の様にして項目ごとに分類されたブロック抽出結果を得ることができる。尚、野線抽出ミス判定処理において、野線抽出ミスがあると判定された場合に、野線抽出ミス判定結果表示処理で、元原稿画像の上に間違っている可能性のある野線を、他の野線と区別しやすく、その野線の色や形状を替えて表示することも可能である。

【0059】尚、上述のステップS208の項目識別処理の処理結果から、更にステップS207のブロック抽

(9)

特開平9-204419

16

出処理のミスの有無を判定し、ブロック抽出処理のミスをオペレータが修正するような処理を更に追加した構成にすることも可能である。例えば、図2のフローチャートのステップS208の項目識別処理の後に、上述の構成を付加した図14に示されるようなフローチャートによって、容易に実現される。

【0060】図14は本実施の形態の他の構成によって実行される処理の処理フローを示すフローチャートである。まず、ステップS1401で、原稿画像をスキャナ101で読み込み2値画像データに変換する。ステップS1402で、得られた2値画像データに対して、2値画像データに含まれる野線を抽出する野線抽出処理を行う。そして、抽出した野線の座標データ等の情報をブロックデータとして図3のように記憶する。

【0061】次に、S1403で、上述のステップS1402の野線抽出処理によって得られた野線に基づいて、野線抽出処理のミスの有無を判定する野線抽出ミス判定処理を行う。ステップS1404で、野線抽出ミス判定処理の処理結果をディスプレイ108に表示する野線抽出ミス判定結果表示処理を行う。ステップS1405で、野線抽出ミス判定結果表示処理の処理結果より、野線抽出処理にミスがあるか否かを判定する。野線抽出処理にミスがない場合、正常終了した旨をディスプレイ108に表示し、ステップS1407に進む。一方、領域分割処理にミスがある場合、領域分割処理にミスがあった旨をディスプレイ108に表示し、領域分割処理のミスの修正をオペレータに促し、ステップS1406に進む。

【0062】ステップS1406で、領域分割判定結果表示処理によって領域分割処理のミスがあった旨が表示された後、元原稿画像上に現在のブロックの領域分割状態が表示される。そして、オペレータは修正対象を拡大、縮小、移動、削除するか、新しいブロックを作成する等の画像処理によって野線抽出処理の処理結果を修正する野線修正処理を行う。

【0063】修正した結果は、オペレータがマウス103で修正終了ボタンを押す等によって指示することで、修正された項目に対応するブロックデータが更新された後、ステップS1407に進む。ステップS1407で、原稿画像に含まれる野線に囲まれた領域（項目）を抽出するブロック抽出処理を行う。そして、抽出した項目の座標データ等の情報をブロックデータとして図3のように記憶する。

【0064】次に、ステップS1408で、原稿画像の項目の構成を示すフォーマット情報からその原稿画像の項目の構成を分類し、その原稿画像を構成する項目を識別する項目識別処理を行う。尚、項目識別処理は、図2のステップS208と同様なので、その詳細は省略する。ステップS1409で、上述のステップS1407のブロック抽出処理によって得られた項目に基づいて、

17

ブロック抽出処理のミスの有無を判定するブロック抽出ミス判定処理を行う。

【0065】ブロック抽出処理にミスがあると判定されるのは、以下に説明されるような場合である。例えば、原稿画像に含まれる文字と罫線との重複やノイズ等によって、図15に示されるような罫線1501が罫線抽出処理によって誤って検出されると、ブロック抽出処理によって本来分離して抽出されるべきでないブロックがブロックB1505とブロックB1503のように分離して抽出されてしまう。この場合、罫線1504等の存在によって、この原稿画像は罫線抽出処理で「筆頭者タイプ」と判定され、罫線抽出ミス判定処理ではミスがないと判定されたとする。そして、ブロック抽出処理において、ブロックB1502は幅が高さよりも長いので「筆頭者タイプ」と判定するが、その下のブロックB1505の幅が高さよりも長いので、「筆頭者タイプ」と識別することに矛盾が生じる。その結果、この場合は、ブロック抽出処理のミスとして判定される。

【0066】ステップS1410で、ブロック抽出ミス判定処理の処理結果をディスプレイ108に表示するブロック抽出ミス判定結果表示処理を行う。ステップS1411で、ブロック抽出ミス判定結果表示処理の処理結果より、ブロック抽出処理にミスがあるか否かを判定する。ブロック抽出処理にミスがない場合、正常終了した旨をディスプレイ108に表示する。一方、ブロック抽出処理にミスがある場合、ブロック抽出処理にミスがあった旨をディスプレイ108に表示し、ブロック抽出処理のミスの修正をオペレータに促し、ステップS1412に進む。

【0067】ステップS1412で、ブロック抽出ミス判定結果表示処理によってブロック抽出処理のミスがあった旨が表示された後、元原稿画像上に現在の項目のブロック抽出状態が表示される。そして、オペレータは修正対象を拡大、縮小、移動、削除するか、新しいブロックを作成する等の画像処理によってブロック抽出処理の処理結果を修正するブロック修正処理を行う。

【0068】修正した結果は、オペレータがマウス103で修正終了ボタンを押す等によって指示することで、修正された項目に対応するブロックデータが更新された後、ステップS1408で、再度、項目識別処理が実行される。例えば、ブロック抽出ミス判定結果表示処理において、ブロック抽出処理にミスがあった旨がディスプレイ108に表示されると、元原稿画像（図6の戸籍文書）の上に処理結果（図15）が表示される。

【0069】元原稿画像画像上に現在の処理結果（図15）が表示されると、オペレータはそこから、間違っ

(10)

特開平9-204419

18

て、修正された項目のブロックの内容は対応するブロックデータに更新され、再度、項目識別処理が実行される。

【0070】そして、項目識別処理の結果からブロック抽出判定処理を行う。この場合、ブロック抽出処理のミスが修正されたので、ブロック抽出判定結果表示処理によってディスプレイ108にその旨が表示された後、正常終了する。また、罫線抽出ミス判定処理で罫線抽出処理のミスが判定された場合に、罫線修正処理において罫線を修正する修正処理と、あるいはそのまま修正せずにブロック抽出処理を行いブロック修正処理において、罫線抽出処理のミスを含むブロックを修正する修正処理のいずれかの修正処理を選択できるような構成にすることで、より処理時間が短くてすみ修正処理を選択することができる。その結果、より効率よく処理全体を実行することができる。

【0071】例えば、図14のフローチャートのステップS1406の後に、上述の構成を付加した図17に示されるようなフローチャートによって、容易に実現される。図17は本実施の形態の他の構成によって実行される処理の処理フローを示すフローチャートである。まず、ステップS1701で、原稿画像をスキャナ101で読み込み2値画像データに変換する。ステップS1702で、得られた2値画像データに対して、2値画像データに含まれる罫線を抽出する罫線抽出処理を行う。そして、抽出した罫線の座標データ等の情報をブロックデータとして図3のように記憶する。

【0072】次に、S1703で、上述のステップS1702の罫線抽出処理によって得られた罫線に基づいて、罫線抽出処理のミスの有無を判定する罫線抽出ミス判定処理を行う。ステップS1704で、罫線抽出ミス判定処理の処理結果をディスプレイ108に表示する罫線抽出ミス判定結果表示処理を行う。ステップS1705で、罫線抽出ミス判定結果表示処理の処理結果より、罫線抽出処理にミスがあるか否かを判定する。罫線抽出処理にミスがない場合、正常終了した旨をディスプレイ108に表示し、ステップS1709に進む。一方、領域分割処理にミスがある場合、領域分割処理にミスがあった旨をディスプレイ108に表示し、ステップS1707に進む。

【0073】ステップS1707で、領域分割処理のミスの修正を行うか否かをオペレータに促す修正選択処理を行う。修正を行う場合、ステップS1708に進む。一方、修正を行わず処理を続行する場合、ステップS1709に進む。ステップS1708で、領域分割判定結果表示処理によって領域分割処理のミスがあった旨が表示された後、元原稿画像上に現在のブロックの領域分割状態が表示される。そして、オペレータは修正対象を拡大、縮小、移動、削除するか、新しいブロックを作成する等の画像処理によって罫線抽出処理の処理結果を修正

19

する罫線修正処理を行う。

【0074】修正した結果は、オペレータがマウス103で修正終了ボタンを押す等によって指示することで、修正された項目に対応するブロックデータが更新された後、ステップS1709に進む。ステップS1709で、原稿画像に含まれる罫線に囲まれた領域（項目）を抽出するブロック抽出処理を行う。そして、抽出した項目の座標データ等の情報をブロックデータとして図3のように記憶する。

【0075】次に、ステップS1711で、原稿画像の項目の構成を示すフォーマット情報からその原稿画像の項目の構成を分類し、その原稿画像を構成する項目を識別する項目識別処理を行う。尚、項目識別処理は、図2のステップS208と同様なので、その詳細は省略する。ステップS1712で、上述のステップS1709のブロック抽出処理によって得られた項目に基づいて、ブロック抽出処理のミスの有無を判定するブロック抽出ミス判定処理を行う。

【0076】ステップS1713で、ブロック抽出ミス判定処理の処理結果をディスプレイ108に表示するブロック抽出ミス判定結果表示処理を行う。ステップS1714で、ブロック抽出ミス判定結果表示処理の処理結果より、ブロック抽出処理にミスがあるか否かを判定する。ブロック抽出処理にミスがない場合、正常終了した旨をディスプレイ108に表示する。一方、ブロック抽出処理にミスがある場合、ブロック抽出処理にミスがあった旨をディスプレイ108に表示し、ブロック抽出処理のミスの修正をオペレータに促し、ステップS1710に進む。

【0077】ステップS1710で、ブロック抽出ミス判定結果表示処理によってブロック抽出処理のミスがあった旨が表示された後、元原稿画像上に現在の項目のブロック抽出状態が表示される。そして、オペレータは修正対象を拡大、縮小、移動、削除するか、新しいブロックを作成する等の画像処理によってブロック抽出処理の処理結果を修正するブロック修正処理を行う。

【0078】修正した結果は、オペレータがマウス103で修正終了ボタンを押す等によって指示することで、修正された項目に対応するブロックデータが更新された後、ステップS1711で、再度、項目識別処理が実行される。そして、項目識別処理の結果からブロック抽出判定処理を行う。この場合、ブロック抽出処理のミスが修正されたので、ブロック抽出判定結果表示処理によってディスプレイ108にその旨が表示された後、正常終了する。

【0079】尚、本発明は、複数の機器（例えばホストコンピュータ、インタフェース機器、リーダー、プリンタ等）から構成されるシステムに適用しても、一つの機器からなる装置（例えば、複写機、ファクシミリ装置等）に適用してもよい。また、本発明の目的は、前述した実

(11)

特開平9-204419

20

施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。

【0080】この場合、記憶媒体から読み出されたプログラムコード自体が上述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。プログラムコードを供給するための記憶媒体としては、例えば、フロッピーディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発性のメモリカード、ROMなどを用いることができる。

【0081】また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているOS（オペレーティングシステム）などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0082】更に、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0083】本発明を上記記憶媒体に適用する場合、その記憶媒体には、先に説明したフローチャートに対応するプログラムコードを格納することになるが、簡単に説明すると、図18あるいは図19のメモリマップ例に示す各モジュールを記憶媒体に格納することになる。すなわち、図18では、少なくとも「記憶モジュール」、「抽出モジュール」、「判定モジュール」および「報知モジュール」の各モジュールのプログラムコードを記憶媒体に格納すればよい。

【0084】尚、「記憶モジュール」は、所定のフォーマットに関するフォーマット情報を記憶する。「抽出モジュール」は、所定のフォーマットを有する原稿画像に含まれる罫線を抽出する。「判定モジュール」は、フォーマット情報に基づいて、抽出結果の正否を判定する。「報知モジュール」は、判定結果に基づいて、報知する。

【0085】また、図19では、少なくとも「記憶モジュール」、「抽出モジュール」、「識別モジュール」、「判定モジュール」および「報知モジュール」の各モジュールのプログラムコードを記憶媒体に格納すればよい。尚、「記憶モジュール」は、複数種類の所定のフォ

(12)

特開平9-204419

21

フォーマットを有する原稿画像に関するそれぞれのフォーマット情報を記憶する。「抽出モジュール」は、所定のフォーマットを有する原稿画像に含まれる罫線を抽出する。「識別モジュール」は、抽出される罫線の特徴に基づいて、所定のフォーマットを有する原稿画像の種類を識別する。「判定モジュール」は、識別された種類に対応する所定のフォーマットを有する原稿画像のフォーマット情報に基づいて、抽出結果の正否を判定する判定。「報知モジュール」は、判定結果に基づいて、報知する。

【0086】

【発明の効果】以上の説明からも明らかなように、本発明によれば、領域分割処理で生じる処理ミスの有無を容易に判断することができ、かつ、処理ミスがある場合にその処理ミスを容易に修正することができる画像処理装置及びその方法を提供できる。また、原稿画像に含まれる罫線を抽出する罫線抽出処理で生じる処理ミスを修正することで、常に正確な罫線を抽出が可能とし、更に、抽出された罫線に基づいて、領域分割処理を行うことで、領域分割処理で生じる処理ミスを軽減する画像処理装置及びその方法を提供することを目的としている。

【0087】また、領域分割処理における操作の作業効率を向上する画像処理装置及びその方法を提供できる。

【図面の簡単な説明】

【図1】本実施の形態の画像処理装置の機能構成を示すブロック図である。

【図2】本実施の形態で実行される処理の処理フローを示すフローチャートである。

【図3】本実施の形態のブロックデータの構成を示す図である。

【図4】本実施の形態の項目識別IDの構成を示す図である。

【図5】本実施の形態のKEYブロックデータ構造体の構成を示す図である。

【図6】本実施の形態の原稿画像の例を示す図である。

【図7】本実施の形態の原稿画像に対して実行した罫線抽出処理の処理結果を示す図である。

【図8】本実施の形態の原稿画像に対して実行した罫線抽出処理にミスがあった時の処理結果を示す図である。

【図9】本実施の形態の原稿画像に対して実行した罫線抽出処理にミスがあった時の処理結果を修正する処理を示す図である。

22

【図10】本実施の形態の項目識別処理の詳細な処理フローを示すフローチャートである。

【図11】本実施の形態のページタイプが「筆頭者タイプ」の場合の項目識別IDを設定する処理の処理フローを示すフローチャートである。

【図12】本実施の形態のページタイプが「非筆頭者タイプ」の場合の項目識別IDを設定する処理の処理フローを示すフローチャートである。

【図13】本実施の形態の入籍データ等の欄の下にあるブロックに項目識別IDを設定する処理の処理フローを示すフローチャートである。

【図14】本実施の形態の他の構成によって実行される処理の処理フローを示すフローチャートである。

【図15】本実施の形態の原稿画像に対して実行したブロック抽出処理にミスがあった時の処理結果を示す図である。

【図16】本実施の形態の原稿画像に対して実行したブロック抽出処理にミスがあった時の処理結果を修正する処理を示す図である。

【図17】本実施の形態の他の構成によって実行される処理の処理フローを示すフローチャートである。

【図18】本発明で実行される処理フローのプログラムを記憶させたFDのメモリマップの構造を示す図である。

【図19】本発明で実行される処理フローのプログラムを記憶させたFDのメモリマップの構造を示す図である。

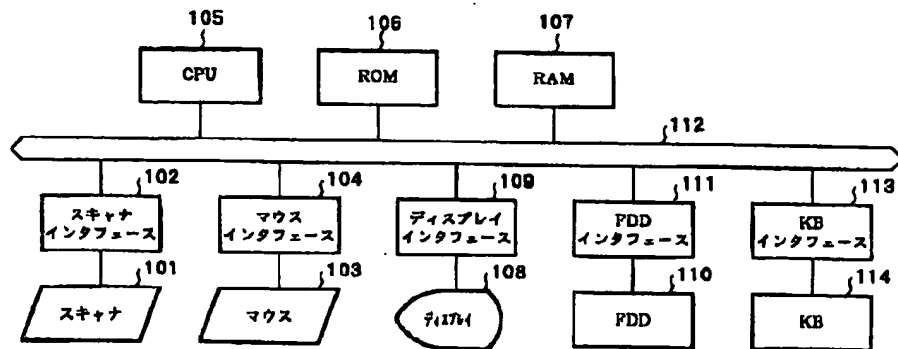
【符号の説明】

- 101 スキャナ
- 102 スキャナインタフェース
- 103 マウス
- 104 マウスインタフェース
- 105 CPU
- 106 ROM
- 107 RAM
- 108 ディスプレイ
- 109 ディスプレイインタフェース
- 110 FDD
- 111 FDDインタフェース
- 112 バス
- 113 KBインタフェース
- 114 KB

(13)

特開平9-204419

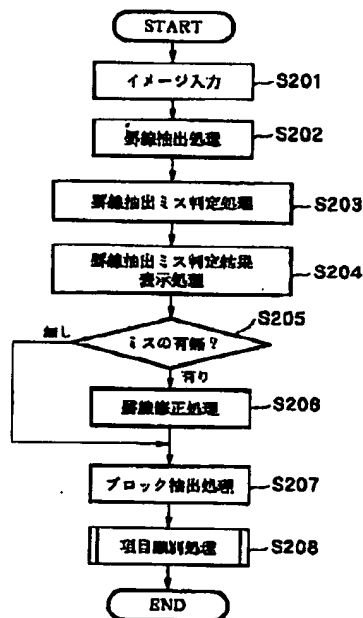
【図1】



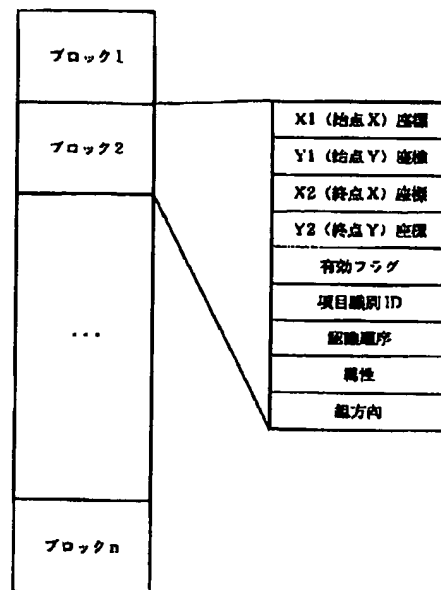
【図18】

ディレクトリ
記憶モジュール
抽出モジュール
判定モジュール
通知モジュール

【図2】



【図3】



【図19】

ディレクトリ
記憶モジュール
抽出モジュール
識別モジュール
判定モジュール
通知モジュール

(14)

特開平9-204419

【図4】

(a)	
フォーマット種別	ID
ページタイプが筆頭者	1000
ページタイプが非筆頭者	2000
(b)	
フォーマット種別	ID
筆頭者のページの記載者	000
非筆頭者のページの第1番目の記載者	100
非筆頭者のページの第2番目の記載者	200
(c)	
項目名	ID
「本籍」というタイトル	10
香地	11
「氏名」というタイトル	12
姓	13
姓事項	20~
「父」というタイトル	30
父名	31
「母」というタイトル	32
母名	33
姓	34
「養父」というタイトル	40
養父名	41
「養母」というタイトル	42
養母名	43
養子続柄	44
「夫」または「妻」というタイトル	50
配偶者名	51
未婚者名	60
「出生」というタイトル	70
出生年月日	71
入籍データ	80~

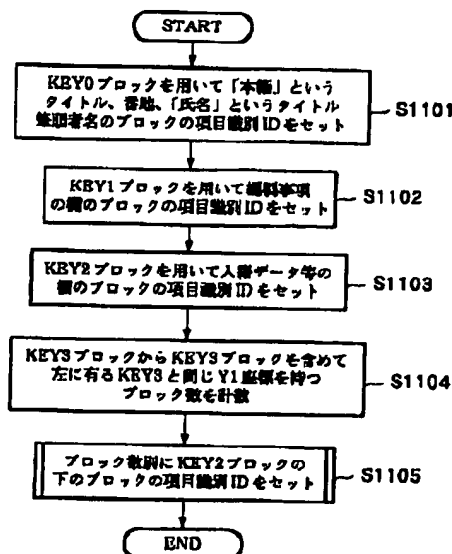
【図5】

KEY0 「本籍」というタイトルがついているブロック
KEY1 出生事項が記載されている欄の最初の行のブロック
KEY2 入籍データ等が記載されている欄の最初の行のブロック
KEY3 「父」というタイトルがついているブロック
KEY4 2人目の入籍データ等が記載されている欄の最初の行のブロック
KEY5 2人目の「父」というタイトルがついているブロック

【図6】

本 籍 東京都大田区下馬三丁目三十番地		氏 名 菅 太郎	
平 成 十 五 年 五 月 六 日 出 生		生 出 平 成 十 五 年 五 月 六 日	
父 名 菅 二 郎		母 名 菅 子 子	
夫 名 菅 太郎		妻 名 菅 子 子	
養父名 菅 二 郎		養母名 菅 子 子	
養子続柄 養子		出生年月日 平 成 十 五 年 五 月 六 日	
入籍データ 平 成 十 五 年 五 月 六 日 入 籍		出生年月日 平 成 十 五 年 五 月 六 日	

【図11】



(15)

特開平9-204419

【図7】

704											701
<p>東京都大田区下丸子三丁目三十番地</p> <p>平成四年四月拾参日編製</p>											
<p>氏名 藤川 太郎</p>											
<p>父 藤川 秀吉 二</p> <p>母 藤川 秀子 男</p> <p>生年 平成四年四月十六日</p>											
709											702
<p>東京都大田区下丸子三丁目三十番地</p> <p>平成四年四月拾参日編製</p>											
<p>氏名 藤川 太郎</p>											
<p>父 藤川 秀吉 二</p> <p>母 藤川 秀子 男</p> <p>生年 平成四年四月十六日</p>											
705											

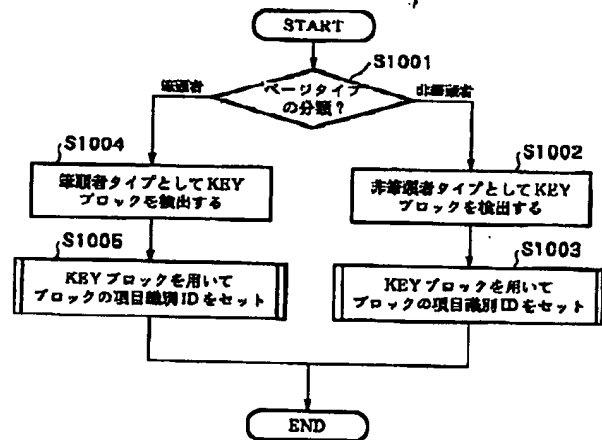
【図8】

804											801
<p>東京都大田区下丸子三丁目三十番地</p> <p>平成四年四月拾参日編製</p>											
<p>氏名 藤川 太郎</p>											
<p>父 藤川 秀吉 二</p> <p>母 藤川 秀子 男</p> <p>生年 平成四年四月十六日</p>											
809											802
<p>東京都大田区下丸子三丁目三十番地</p> <p>平成四年四月拾参日編製</p>											
<p>氏名 藤川 太郎</p>											
<p>父 藤川 秀吉 二</p> <p>母 藤川 秀子 男</p> <p>生年 平成四年四月十六日</p>											
805											

【図9】

修正終了		902
<p>東京都大田区下丸子三丁目三十番地</p> <p>平成四年四月拾参日編製</p>		901
氏名		

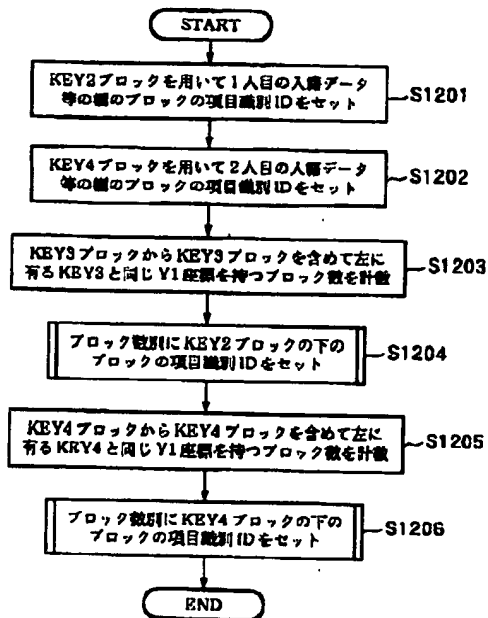
【図10】



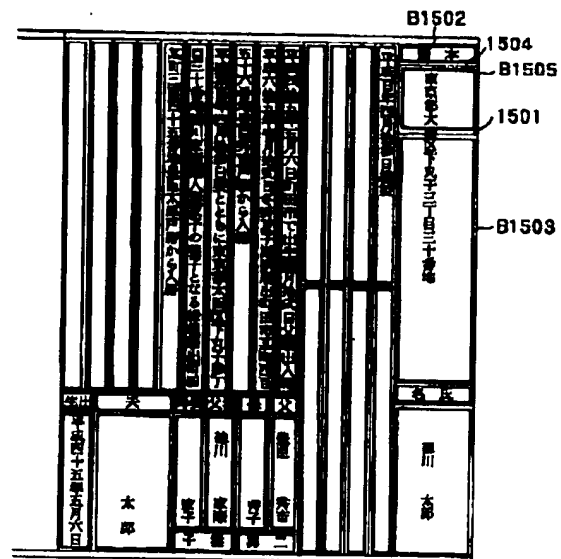
(16)

特開平9-204419

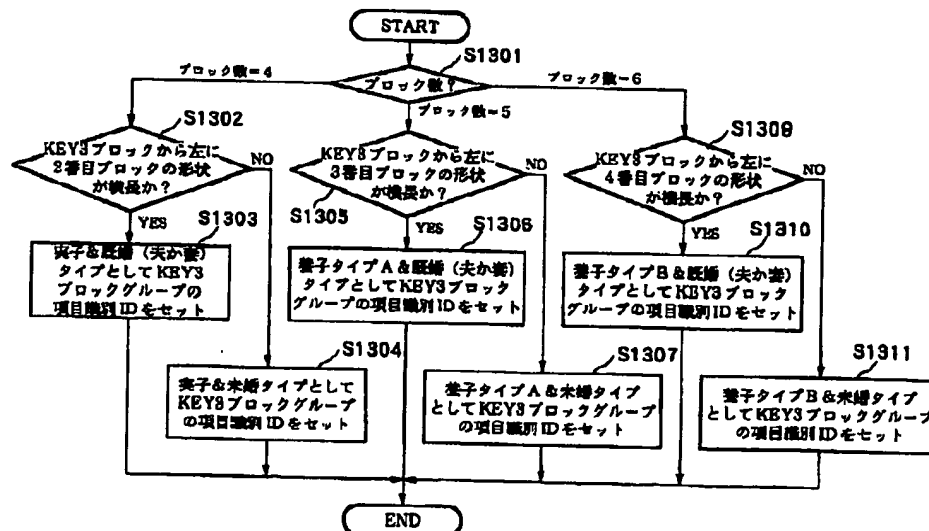
【図12】



【図15】



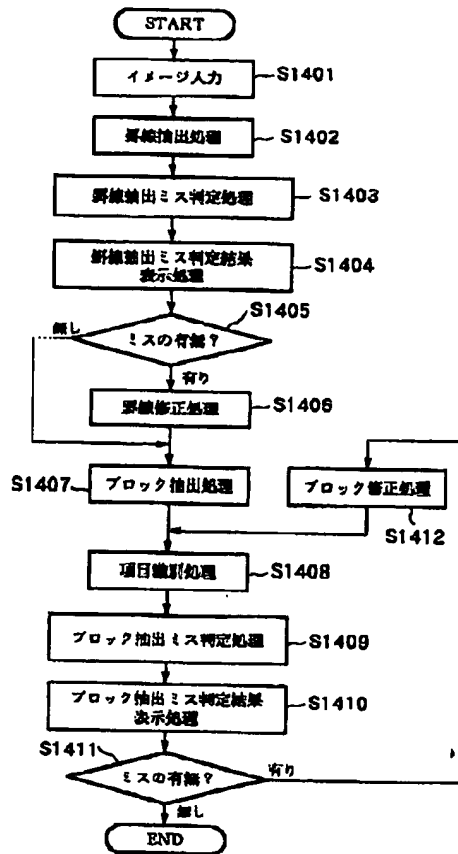
【図13】



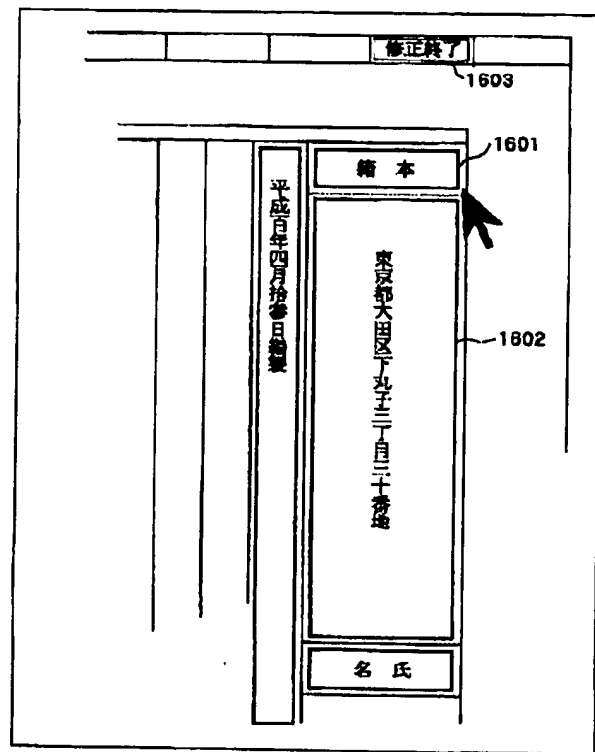
(17)

特開平9-204419

【図14】



【図16】



(18)

特開平9-204419

【図17】

